

Pixel-wise Land-Cover Classification Techniques of Satellite Optical Images

Report attività primo anno
Dottorato di ricerca in Fisica XXXIII ciclo

Dottorando: Roberto Cilli
Tutori: prof. Roberto Bellotti, dott. Antonio Zilli

January 22, 2019

In questo report sarà esposta l'attività di ricerca portata avanti dal candidato durante il primo anno di Dottorato presso il Dipartimento di Fisica dell'Università degli Studi di Bari. L'attività di ricerca del candidato è stata incentrata sulla implementazione di tecniche di Pattern Recognition per l'analisi e l'elaborazione di immagini satellitari ottiche provenienti dalla missione Sentinel-2 con l'obiettivo di rivelare automaticamente presenza di nuvole (Cloud Detection) o uso del suolo (LULC Classification) in una certa regione di interesse (ROI). Tale attività di ricerca, dal carattere fortemente applicativo, può avere potenziale impatto su temi quali ad esempio monitoraggio ambientale, gestione disastri ed emergenze umanitarie, censimenti in aree urbane e gestione della rete idrica urbana. In seguito, sarà presentato lo stato dell'arte dell'attività di Dottorato.

Programma Copernicus e Missione Sentinel-2

Il Programma Europeo di osservazione della Terra Copernicus, precedentemente conosciuto come GMES (Global Monitoring for Environment and Security), è un insieme complesso di sistemi che raccoglie informazioni da molteplici fonti, ossia satelliti di osservazione della Terra, sensori di terra e sensori aviotrasportati. Integra ed elabora tutte queste informazioni, fornendo agli utenti, istituzionali ed afferenti al comparto industria, informazioni affidabili ed aggiornate attraverso una servizi che attengono all'ambiente, al territorio ed alla sicurezza.

Copernicus ha tra i suoi obiettivi quello di garantire all'Europa una sostanziale indipendenza nel rilevamento e nella gestione dei dati di monitoraggio Terrestre. A tale scopo, l'agenzia spaziale europea (ESA) sta sviluppando sei famiglie di missioni satellitari, note come missioni Sentinel, progettate per le esigenze operative del programma Copernicus. L'attività di ricerca del dottorando ha previsto l'analisi delle immagini satellitari provenienti dalla missione Sentinel-2.

Sentinel-2 è una missione ottica multispettrale costituita da due satelliti identici, Sentinel-2A e -2B, che operano sulla stessa orbita elio-sincrona e sfalsati di 180 gradi per garantire l'osservazione sistematica delle terre emerse con lo stesso angolo visuale ed alla stessa ora locale, con un intervallo di tempo tra un'acquisizione e la successiva di 5 giorni. Ciascun satellite è equipaggiato con 13 sensori ottici (B1-B12) in grado di coprire un intervallo di lunghezze d'onda dello spettro e.m. dal visibile all'infrarosso (SWIR, short-wave infrared) con risoluzione spaziale di 10, 20 o 60 metri. In figura 1 sono riportate lunghezza d'onda centrale, larghezza di banda e risoluzione spaziale di ciascuno dei 13 sensori ottici montati sui satelliti Sentinel-2.

Le acquisizioni della superficie Terrestre realizzate dalla missione Sentinel-2 sono disponibili liberamente e distribuiti come:

- prodotti L1B, ovvero immagini di radianza top-of-atmosphere (TOA) che richiedono esperti di dominio in ortorettificazione delle immagini;
- prodotti L1C, ovvero immagini di riflettività TOA georeferenziate in coordinate UTM-WGS84. Tali immagini, relative a regioni di area $110 \text{ Km} \times 110 \text{ Km}$, presentano correzioni radiometriche e geometriche (ortorettificazioni incluse). La definizione di riflettività è discussa in appendice;
- prodotti L2A, immagini di riflettività bottom-of-atmosphere (BOA) in coordinate UTM-WGS84. Tali immagini, presentano correzioni radiomet-

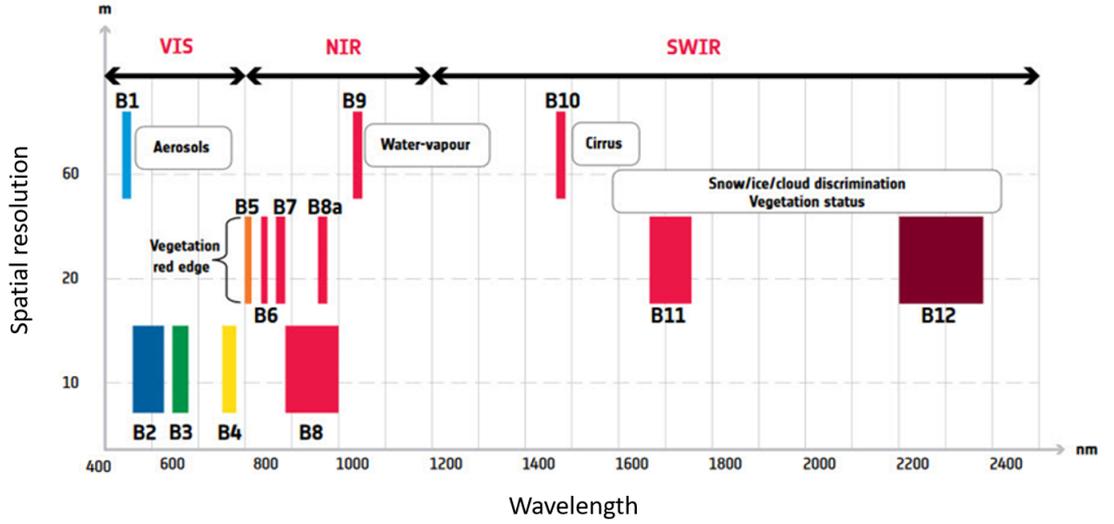


Figure 1: Specifiche dei sensori montati a bordo dei satelliti Sentinel-2.

riche, geometriche ed atmosferiche. Sono realizzate tramite software prototipali distribuiti da ESA (Sen2Cor) o da CNES/CESBIO (MAJA).

L'attività di Dottorato ha previsto, durante l'anno, un uso intensivo delle immagini Sentinel-2 L1C al fine di sviluppare una pipeline di elaborazione delle immagini finalizzata alla rivelazione automatica di nuvole e di classificazione di regioni di interesse (ROI) in base all'uso/copertura del suolo, mediante tecniche di machine learning come support vector machine, random forest e multi-layer perceptron.

Calcolo della riflettività TOA delle immagini Sentinel-2 L1C

I sensori a bordo dei satelliti Sentinel-2 misurano radianze TOA $R_k(x, y)$ per ciascun sensore k e per ciascun pixel (x, y) . La conversione delle radianze TOA in riflettività TOA $\rho_k(x, y)$ si basa sul modello di irradianza solare e sull'angolo Sole-Zenit:

$$\rho_k(x, y) = \frac{\pi \times R_k(x, y)}{A_k \times E_{S,k} \times \cos(\theta_{SZ}(x, y))} \times d^2(t) \quad (1)$$

in cui:

- A_k è un coefficiente di calibrazione del sensore k -esimo;
- $E_{S,k} = \int d\lambda T_k(\lambda) R_S(\lambda)$ è lo spettro solare extraterrestre pesato sulla funzione di risposta del sensore k -esimo $T_k(\lambda)$;
- $\theta_{SZ}(x, y)$ è l'angolo compreso tra Zenit e direzione Terra-Sole;
- $d(t) = d_{TS} \times (1 - \epsilon \times \cos(\omega t))$ è la distanza Terra-Sole al tempo t di acquisizione della superficie Terrestre.

Le immagini Sentinel-2 L1C sono immagini di riflettività TOA, una immagine per sensore, codificate come numeri interi a 15 bit. Il valore delle riflettività varia da 1 (minima riflettività) a 10000 (riflettività 1), ma valori più alti possono essere osservati in alcuni casi. Il valore 0 è riservato per le regioni non acquisite (no data).

Dataset per Segmentazione semantica di ROI in immagini satellitari

Il seguente paragrafo introduce i dataset analizzati ed i problemi affrontati durante il primo anno di dottorato:

- rivelazione di nuvole con dataset pugliese;
- rivelazione uso del suolo (LULC classification) con dataset Hollstein et al.;
- rivelazione di edifici in immagini satellitari ad altissima risoluzione di aree urbane densamente popolate.

Dataset pugliese e rivelazione nuvole

Le perturbazioni atmosferiche come nuvole e cirri coprono oltre il 50% della superficie Terrestre, giocano un ruolo importante nel bilancio energetico del pianeta a causa del loro ruolo nell'assorbimento e nello scattering della radiazione e.m. solare. La presenza di nuvole, come è noto alla comunità di remote sensing, è altamente indesiderata in tutte quelle situazioni in cui è necessario osservare fenomeni sulla superficie Terrestre. In molte situazioni, è necessario acquisire immagini prive di nuvole o cirri, dato che la loro presenza costituisce rumore e rende le misure di radianza riflessa TOA non attendibili. L'informazione relativa a regioni coperte da nuvole è dunque persa.

L'eliminazione efficace di regioni coperte da nuvole e cirri risulta pertanto una priorità per la comunità di remote sensing. Differenti algoritmi basati su test empirici sono stati proposti. A titolo di esempio si possono citare gli algoritmi MAJA sviluppato da CNES e CESBIO, Sen2Cor commissionato da ESA ed FMask sviluppato da Zhu et al.

Per affrontare il problema della rivelazione delle regioni nuvolose in immagini satellitari ottiche, sono stati sviluppati diversi algoritmi di apprendimento supervisionato basati sui classificatori binari descritti nella sezione precedente. Gli algoritmi sono stati allenati e testati su un dataset di immagini Sentinel-2 L1C relative all'area Nord-Barese e Murgia (Dataset pugliese). In seguito si riportano le caratteristiche del dataset adoperato:

- 155 immagini Sentinel-2 L1C della medesima regione acquisite a tempi distinti;
- immagini di area $1 \text{ Km} \times 1 \text{ Km}$ (100×100 pixel);
- immagini munite di etichette binarie (0 cielo sereno, 1 cielo nuvoloso);

Dataset Hollstein et al. e LULC classification

La classificazione di nuvole, cirri, nuvole, ombre, neve e regioni di cielo sereno in immagini satellitari ottiche (LULC classification) è uno step cruciale nella pre-elaborazione delle immagini satellitari e può essere utile per eseguire correzioni atmosferiche. Il problema di LULC classification di immagini satellitari si può ritenere una generalizzazione del problema della rivelazione delle nuvole.

Tale problema è stato affrontato seguendo un approccio supervisionato basato sulle tecniche di pattern recognition introdotte nel paragrafo precedente. A tal fine, è stato impiegato un dataset etichettato (dataset Hollstein et al.) per addestrare e testare diversi algoritmi di apprendimento supervisionato. Ogni pixel del dataset Hollstein et al. è etichettato secondo sei classi semantiche:

- clear-sky (cielo sereno)
- water (fiumi, laghi, oceani)
- shadow (ombre di nuvole e topologiche)
- cirrus (cirri e nuvole sottili)
- cloud (nuvole spesse)
- snow (neve o ghiaccio)

Il dataset Hollstein et al., diversamente dal dataset pugliese, è costituito da porzioni di immagini Sentinel-2 L1C campionate casualmente e provenienti da oltre 30 paesi, per un totale di oltre 5 milioni di pixel. Questo rende il dataset per LULC classification più eterogeneo, quindi maggiormente rappresentativo delle possibili aree della superficie Terrestre.

Dataset Fiumicino 2015, Fiumicino 2017 e Venezia 2018 per Building Detection

La classificazione di edifici in immagini satellitari ad alta risoluzione è un compito le cui finalità possono essere quelle di censimento della popolazione o stime del PIL di una certa area, valutazione della area disponibile per l'installazione di pannelli fotovoltaici (calcolando l'area disponibile sui tetti ad esempio) o la realizzazione di mappe urbane. Tuttavia la segmentazione manuale di immagini satellitari di aree urbane è una operazione ripetitiva e che richiede tempo. Pertanto la segmentazione automatica di palazzi è un argomento molto caldo sia per la comunità di remote sensing sia per la comunità di computer-vision.

Dato il crescente interesse per la rivelazione automatica di palazzi (Building Detection), sono stati impiegati metodi di machine learning tradizionali su tre dataset etichettati di immagini satellitari ad altissima risoluzione (very high resolution, VHR) le cui proprietà brevemente riassunte:

- 1,2. dataset Fiumicino 2015 e 2017: immagini 10160×8400 pixel, RGB, VHR, georeferenziate, della regione di Fiumicino acquisite a due anni distinti dai satelliti Pleiades della agenzia spaziale francese (Centre National Etudies Spatiale, CNES). Risoluzione spaziale: 50×50 cm²;

3. dataset Venezia 2018: immagine VHR, georeferenziata, 8185×7935 pixel, 8 bande multispettrali ottiche, dell'area metropolitana di Venezia acquisita dal satellite commerciale WorldView-3. Risoluzione spaziale: 40×40 cm²;

Le etichette associate ai 3 dataset sono fornite da Open Street Map, un progetto collaborativo finalizzato a creare mappe a contenuto libero del mondo. Le etichette sono realizzate da utenti esperti di dominio su base volontaria. Ad ogni pixel è associata una etichetta binaria: 1 per pixel di tipo "building/built-up area", 0 per tutti gli altri pixel.

Pipeline di elaborazione delle immagini satellitari

Il presente paragrafo si propone di descrivere la pipeline di classificazione automatica delle immagini Sentinel-2 realizzata durante l'attività di dottorato e schematizzata in figura 2. L'approccio seguito al fine di classificare le immagini è stato quello di un apprendimento supervisionato basato sui classificatori precedentemente descritti (SVM, MLP e RF). A ciascun pixel appartenente alle immagini analizzate è associata una etichetta, detta anche classe o verità di base, derivante da un'analisi visuale effettuata da un esperto di dominio. L'idea alla base della seguente pipeline di analisi è quella di fornire un numero di esempi per ogni classe semantica sufficiente per far in modo che l'algoritmo sviluppato sia in grado di classificare correttamente esempi mai visti prima dall'algoritmo.

Come schematizzato in figura 2, l'intero dataset di immagini a disposizione viene suddiviso in due sottoinsiemi, detti *training set* e *test set*. Il primo insieme, viene impiegato per addestrare l'algoritmo a classificare i pixel di una certa immagine multispettrale secondo alcune classi semantiche definite a priori in base al problema che si intende affrontare. Il secondo insieme, invece, viene adoperato per verificare che l'algoritmo sviluppato sia effettivamente in grado di generalizzare la classificazione per immagini mai viste prima dall'algoritmo stesso. La suddivisione delle immagini del dataset è ripetuta diverse volte ed in maniera casuale (random splitting). Tale suddivisione casuale del dataset, nota in letteratura come ciclo di cross-validation, è ripetuta 100 volte.

Ad ogni ciclo di cross validation, viene eseguito un campionamento casuale di un numero fissato di pixel (da 100 a 1,000,000 pixel) appartenenti alle immagini del training set. I pixel campionati con le etichette relative sono impiegati come esempi per addestrare un classificatore supervisionato. Una volta addestrato il classificatore, questo viene testato sul test set, e le performance di classificazione vengono valutate secondo alcune metriche, quali accuracy oppure overall accuracy.

Il campionamento casuale dei pixel è ripetuto per 50 volte ad ogni ciclo di cross-validation. Le performance del classificatore sul test set, corrispondenti ad ogni campionamento, vengono calcolate e se ne determina la media. Tale media costituisce una stima delle prestazioni del classificatore ad un ciclo di cross-validation fissato. Infine, si determina media e deviazione standard delle (medie delle) prestazioni calcolate ad ogni ciclo di cross-validation, ottenendo in tal modo una stima oggettiva della capacità di classificazione dell'algoritmo con un errore associato.

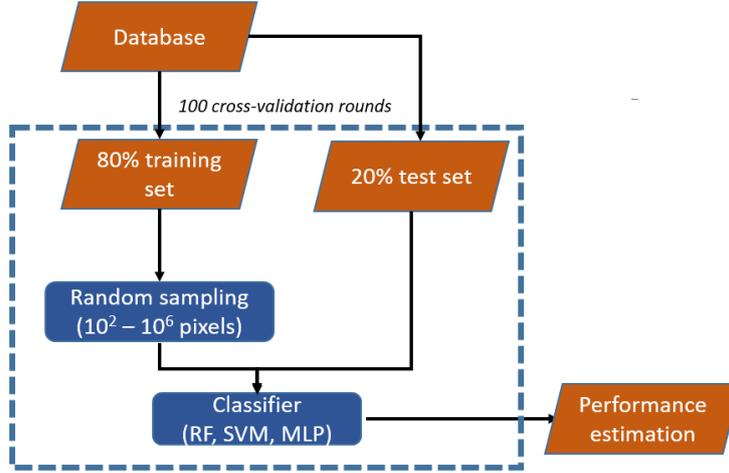


Figure 2: Pipeline per la classificazione automatica dei pixel di una immagine Sentinel-2. Tale pipeline è stata adoperata per i problemi di Cloud Detection e LULC classification. Il codice sorgente della pipeline di analisi è stato scritto in linguaggio R.

Risultati

Il presente paragrafo si occuperà di introdurre i risultati ottenuti sui due dataset precedentemente introdotti. La pipeline del paragrafo precedente è stata implementata su entrambi i dataset e sono state valutate le prestazioni al variare del numero di pixel casualmente campionati ed al variare dei classificatori riportati in tabella 1.

Rivelazione di nuvole

Le prestazioni della pipeline per il problema della rivelazione delle nuvole sono state valutate mediante le metriche *accuracy*, *sensitivity* e *specificity*, definite come segue:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$SEN = \frac{TP}{TP + FN} \quad (3)$$

$$SPE = \frac{TN}{TN + FP} \quad (4)$$

I classificatori sono stati addestrati come descritto nel paragrafo precedente e le performance di classificazione valutate mediante *accuracy* al variare del numero dei pixel campionati casualmente. In figura 3 sono riportati i boxplot relativi alle accuracy misurate sugli insiemi di test e di training usando random forest. Si può notare come il numero di esempi necessario per ottenere le massime performance di classificazione possibili è dell'ordine di 10^3 indipendentemente dal classificatore impiegato. Tale saturazione delle performance dei classificatori è riconducibile alla elevata correlazione dei pixel che costituiscono

modello	parametri
RF	ntree = 500 nsplit = 4
SVM	kernel = "radial" $\gamma = 1/14$ C = 1 $\mu = 0.5$ $\epsilon = 0.1$ tolleranza = 0.001
MLP	neuroni nascosti = 5 activation function = "sigmoid" learning rate = 0.1 momentum = 0.05

Table 1: Sommario dei classificatori adoperati per la implementazione della pipeline schematizzata in figura 2

il dataset pugliese (si pensi a pixel adiacenti o pixel relativi ad uno stesso punto acquisito diverse volte).

La procedura appena descritta è stata ripetuta per ogni classificatore riportato in tabella 1. In tabella 2 sono riportate le performance dei classificatori proposti e dell'attuale stato dell'arte nella rivelazione automatica delle nuvole rappresentato dall' algoritmo MAJA.

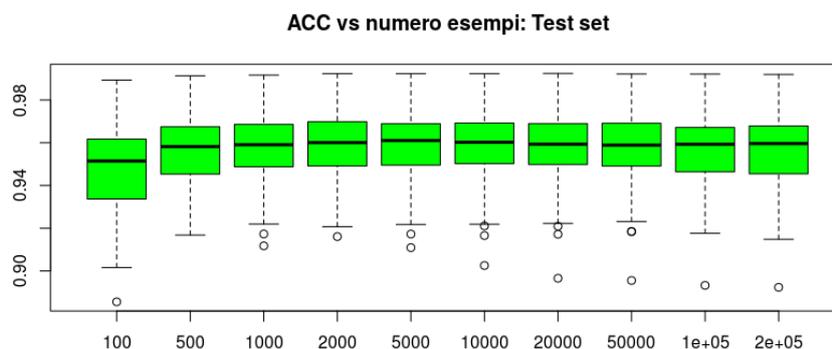


Figure 3: Boxplot della accuratezza della pipeline in funzione del numero di pixel casualmente campionati dal training set, per il problema della rivelazione delle nuvole in immagini satellitari. Classificatore impiegato: Random Forest.

LULC classification

Le prestazioni dell' algoritmo di classificazione dell'uso e della copertura del suolo (LULC classification) sono state espresse mediante *overall accuracy* (OA),

	RF	SVM	MLP	MAJA
ACC	96 ± 2	96 ± 3	87 ± 15	93 ± 3
SEN	93 ± 4	90 ± 8	92 ± 7	99.3 ± 0.6
SPE	97 ± 2	98 ± 2	85 ± 25	90 ± 5

Table 2: Confronto delle prestazioni dei classificatori in tabella 1 e confronto con lo stato dell’arte nella rivelazione delle nuvole mediante le metriche accuracy, sensitivity e specificity.

	OA (%)
RF	88 ± 1
SVM	91 ± 3
MLP	88 ± 17

Table 3: Confronto delle prestazioni dei classificatori in tabella 1 espresse mediante OA per il problema di LULC classification sul dataset Hollstein et al.

definita come:

$$OA = \frac{\text{\#esempi correttamente etichettati}}{\text{\#esempi totali}} \quad (5)$$

Tale metrica è una generalizzazione dell’accuracy per un problema a più di due classi.

Come per il problema della rivelazione delle nubi, la pipeline è stata adoperata per i classificatori elencati in tabella 1 e per diversi numeri di pixel campionati casualmente. In figura 4, sono mostrate le OA su test set ottenute addestrando SVM su diversi numeri di pixel campionati e la matrice di confusione, su test set, ottenuta addestrando SVM con un numero di pixel campionati pari ad 1 milione.

Building Detection

Le prestazioni dei classificatori introdotti in 1, espresse in funzione delle metriche ACC, SEN, e SPE introdotte in (2),(3) e (4), sono riportate in tabella 4. È possibile notare che i classificatori SVM e RF hanno performance identiche sui dataset Fiumicino 2015 e Venezia 2018 (ACC su problema binario inferiore a 80%). Tutti i dataset impiegati per Building Detection sono sbilanciati, ovvero il numero di pixel di tipo building è circa il 7 % dei pixel totali costituenti il dataset. Per evitare che pochi esempi di tipo Building siano presenti durante la fase di addestramento dei classificatori, è stato eseguito un campionamento casuale dei pixel del training set imponendo un numero di esempi di tipo Building e non-Building uguale (ovvero è stato imposto un rapporto unitario del numero di esempi delle due classi nel training set). In questo modo è stato possibile massimizzare la sensitivity dei classificatori RF ed SVM a circa 75 % per il dataset Fiumicino 2015 e ad 82 % per il dataset Venezia 2018. Le prestazioni ottenute con i metodi di classificazione tradizionali non sono sufficientemente elevate per poter pensare di impiegare tali classificatori per poter realizzare mappe di regioni urbane nè per poter rivelare con sufficiente sicurezza cambiamenti nell’uso del suolo mediante confronto di coppie di immagini della

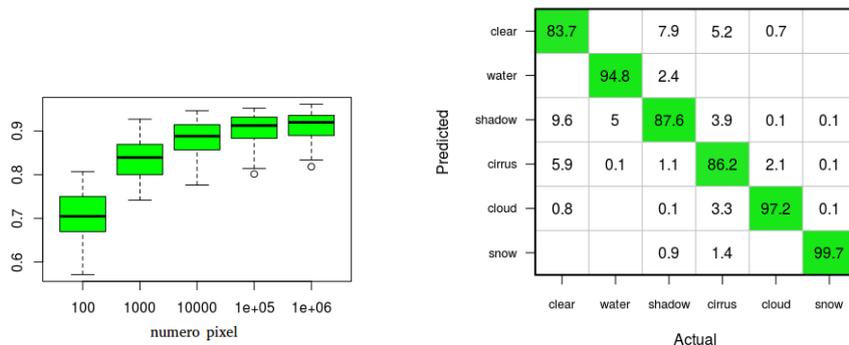


Figure 4: A sinistra: Boxplot di OA in funzione del numero di pixel casualmente campionati dal training set, per il problema di LULC classification sul dataset Hollstein et al. Classificatore impiegato: Support Vector Machine. A destra: Matrice di confusione relativa ottenuta addestrando SVM con 1 milione di pixel casualmente campionati. Si può notare che le classi di pixel più frequentemente fraintese siano "clear", "cirrus" e "shadow".

	Fiumicino 2015		Venezia 2018	
	RF	SVM	RF	SVM
ACC	78 ± 10	78 ± 10		80 ± 20
SEN	75 ± 14	75 ± 12		80 ± 20
SPE	78 ± 12	78 ± 11		75 ± 20

Table 4: Confronto delle prestazioni dei classificatori in tabella 1 mediante le metriche accuracy, sensitivity e specificity per il problema di Building Detection (Dataset Fiumicino 2015 e Venezia 2018, numero di pixel usati in training 20,000).

stessa regione acquisita più volte (Change Detection o rivelazione di nuovi edifici). La seguente analisi è stata discussa mediante poster in occasione della ESA Earth Observation ϕ -week tenutasi dal 12 al 16 Novembre 2018.

Attività di ricerca futura

Gli sviluppi futuri della attività di ricerca nel prossimo anno prevedono uno studio accurato dello stato dell'arte delle tecniche di machine learning tradizionali applicate alla elaborazione di segnali da telerilevamento con l'obiettivo di scrivere un paper circa l'analisi del dataset Hollstein et al. per LULC classification. Inoltre, per quanto concerne il problema della rivelazione dei palazzi, si cercherà di sviluppare un algoritmo di segmentazione semantica delle immagini satellitari basato su una rete neurale convoluzionale (basata su U-net) che possa rivelare con più efficacia gli edifici in immagini VHR di aree urbane densamente popolate con l'obiettivo di sviluppare un software applicativo per la rivelazione di nuovi edifici (Change Detection) mediante analisi di coppie di immagini satellitari della medesima regione acquisita ad anni differenti.

Corsi seguiti

1. "How to prepare a technical speech in English", Prof.ssa White
2. "Management and knowledge of european research and promotion of research results", Prof.ssa D'Orazio
3. "Programming with Python for Data Science", Prof.Diacono
4. "Elaborazione Segnali ed Immagini", Prof. Bellotti
5. "Introduction to parallel Computing and GPU Programming using CUDA, Dott. Pantaleo
6. "Interpolation Methods and Techniques for Experimental Data Analysis", Prof. Pompili
- X. "Introduction to C++ Programming", Prof. Cafagna

Conferenze

1. The ESA Earth Observation ϕ -week Open Science and Future EO, Frascati, ESRIN, 12-16/11/2018.

Seminari

1. "Cycle of lectures on SM and BSM models". Bari, Dipartimento Interateneo di Fisica, Prof. S Khalil - Director of the Center for Theoretical Physics Zewail City for Science and Technology, Egypt. 20-23/03/2018
2. "The Evolution of Countries' Production Specialization and the Design of Industrial Policies", Prof. Coniglio - Dipartimento Economia e Finanza, Università di Bari. 8/10/2018
3. "United Nations eGovernement Survey 2018: Gearing eGovernement to Support Transformation toward Sustainable and Resilient Societies", Ing. Aquaro - Direttore del Digital Governement Branch presso il Dipartimento degli Affari Economici e Sociali delle Nazioni Unite. 8/10/2018